

**Request to Archive
With The National Centers for Environmental Information
For GOES-R Level 2+ Space Weather Data
Provided by NOAA/NWS**

2016-04-08

This information will be used by NCEI to conduct an appraisal and make a decision on the request.

1. Who is the primary point of contact for this request?

Meg Tilton

NCEI

Data Manager

303-497-6223

margaret.tilton@noaa.gov

Normal working hours: Monday 1 - 5 p.m, Tues-Thur. 9 a.m. - 5 p.m., Friday 9 a.m. - 1 p.m.

2. Name the organization or group responsible for creating the dataset.

National Weather Service

3. Provide an overview summarizing the scope of data you want to archive. Describe the outputs, data variables, including their measurement resolution and coverage.

NCEI plans to archive all L2+ GOES-R Space Weather data produced by the National Weather Service. This includes the following products:

1-min averages for XRS, EUVS, magnetic field, and the SEISS's Solar Galactic Proton Sensor (SGPS) and Magnetospheric Particle Sensors - High and Low (MPSH and MPSL)

1-sec irradiance flux for XRS

XRS and EUVS event detection

XRS post-event summary

XRS flare location

EUVS daily averages

XRS daily background

Magnetic field measurements in additional coordinate systems

Quiet magnetic fields

Magnetopause crossing

5-min averages for SEISS's MPSH, MPSL, and SGPS

SGPS integral flux

SGPS event detection

MPSH and MPSL moments

SGPS integral flux

SGPS rate of rise

Linear energy transfer from the SEISS's Energetic Heavy Ion Sensor (EHIS)

Data from the Solar Ultraviolet Imager (SUVI): composite images, thematic maps, bright regions, flare locations, coronal holes and coronal hole boundaries, running difference images, and fixed difference images

4. What is the time period covered by the dataset? (YYYY-MM-DD, YYYY-MM or YYYY)

From 2016-10-13

Ongoing as continuous updates to the data record

5. Edition or version number(s) of the dataset:

1.0.0

6. Approximate date when the dataset was or will be released to the public:

2017-04-13

7. Who are the expected users of the archived data? How will the archived data be used?

Space weather forecasters and modelers, including the Space Weather Prediction Center (SWPC); scientists at academic, research and government institutions

8. Has the dataset undergone user evaluation and/or an independent review process? Did NCEI participate in design reviews?

NCEI is developing a demonstration software system to produce these products. This system will then be turned over to the NWS, which will operate their own instance of it in order to produce operational L2+ products used in forecasting. During the development of the demonstration system, NCEI consulted extensively with the Space Weather Prediction Center (SWPC) to ensure L2+ products met their forecasting needs.

9. Describe the dataset's relationship to other archived datasets, such as earlier versions or related source data. If this is a new version, how does it improve upon the previous version(s)?

L2+ GOES-R products are higher-level products derived from GOES-R L1b data provided by the Ground System. Compared to L1b products, these higher-level products facilitate forecasting of space weather events. L1b data are in turn generated by processing L0 data, which NCEI is also archiving for Space Weather products.

10. List the input datasets and ancillary information used to produce the data.

Input datasets: GOES-R L1b data. These data are produced by the GOES-R Ground System until handover, and afterwards by the Office of Satellite Products and Operations (OSPO).

Real-time Ancillary Data: Real-time Solar Wind & Inter-planetary Magnetic Field Data; GOES-NOP Magnetic Field Data

Quasi-static Ancillary Data: International Magnetic Field Model (IGRF); Olson Pfitzer Magnetic Field Model (OP77); Solar Beta and P angles, Solar Distance from USNO; USAF/NOAA Solar Region Summaries (daily)

11. List web pages and other links that provide information on the data.

<https://drive.google.com/drive/search?q=SPADES%20CDR>

<http://www.goes-r.gov/>

http://www.ngdc.noaa.gov/stp/space-weather/online-publications/stp_sii/spades/

12. List the kinds of documents, metadata and code that are available for archiving. For example, data format specifications, user guides, algorithm documentation, metadata compliant with a standard such as ISO 19115, source code, platform/instrument metadata, data/process flow diagrams, etc.

1. netCDF-4 files for data derived from the following instruments: Extreme Ultraviolet and X-ray Irradiance Sensor (EXIS), Space Environment In-Situ Suite (SEISS), and magnetometer (MAG). Some quality data for the Solar Ultraviolet Imager (SUVI) is also contained in netCDFs.
2. FITS files derived from data produced by the Solar Ultraviolet Imager (SUVI).

3. Python code (.py files) for implementing scientific algorithms to create L2+ products.
4. XML files containing collection-level metadata in XML format.
5. Additional documentation in the form of data notes, peer-reviewed articles, algorithm dependency diagram, and so on.

13. Indicate the data file format(s).

1. netCDF-4
2. FITS
3. XML
4. PDF

14. Are the data files compressed?

gzip

15. Provide details on how the files are named and how they are organized (e.g., file_name_pattern_YYYYMM.tar in monthly aggregations).

L2+ files include both daily files and files generated at a shorter cadence.

Daily file names will have the format Env_DSN_PID_Date_Version.Ext, as follows:

Env = Environment

DSN = Data Short Name (e.g. mpsl-l2-avg1m)

PID = Platform ID (e.g. g16 for GOES-16 / geo for GOES non-specific)

Date = Product Start Date (dYYYYMMDD)

Version = Version = Version Number (va_b_c) (revision: a - major / b - minor / c - patch)

Ext = Extension (e.g., "nc" for netCDF4, "fits" for FITS; note FITS files will likely be tarred)

Granular file names will have the format Env_DSN_PID_Start_End_Version.Ext. Fields are the same as above except for Start and End, which have the format sYYYYMMDDhhmmss and eYYYYMMDDhhmmss.

Details on the possible Environment and DSN values are given in a Google doc:

<https://docs.google.com/spreadsheets/d/1lwsQZz3IlCR5kymqKv6fDFtnQa5kUqiRSibbSrH8gJo/edit#gid=0>

16. Explain how to access sample data files and/or a file listing for previewing. If it is not available now, when will it be available?

Sample files are in the following Google drive folder:

<https://drive.google.com/a/noaa.gov/folderview?id=0B7prHmLw9LYmZ0FHVXBoQUPEd2s&usp=sharing>

Note file header metadata will likely change.

17. What is the total data volume to be submitted?

Continuous Data: data volume rate for a continuous data production.

Total Data Volume Rate: 40TB per Year

Data File Frequency: 68 per Day

Data Production Start: 2017-01-02

18. Are later updates, revisions or replacement files anticipated? If so, explain the conditions for submitting these additional data to the archive.

The file names contain a version number, so any reprocessed files will be stored in addition to the originals.

19. Describe the server that will connect to the ingest server at NCEI for submitting the data.

Physical Location: Boulder, CO; Backup is College Park, MD (NCEP)
System Name: Integrated Dissemination Program (IDP)
System Owner: DOC/NOAA/NWS > National Weather Service, NOAA, U.S.
Department of Commerce

Additional Information:

20. What are the possible methods for submitting the data to NCEI? Select all that apply.

Currently unknown. Data will be ingested from NWS's Integrated Dissemination Program (IDP) system.

21. Identify how you would like NCEI to distribute the data. Web access support depends on the resources available for the dataset.

1. User interface to order and stage data for download
2. Direct download links

22. Will there be any distribution, usage, or other restrictions that apply to the data in the archive?

Constraint Type	Description
Access	Access to Post-Launch Test (PLT) data is restricted during the actual PLT period.

23. Discuss the rationale for archiving the dataset and the anticipated benefits. Mention any risks associated with not archiving the dataset at NCEI.

GOES-R is a data set of considerable importance. The L2+ data will offer key indicators of solar storms and will be used in SWPC's forecasts. Archiving the operational data produced by NWS will allow scientists to refine their space weather forecasting models. Failure to archive these data would result in an interruption to a long-standing data set, as GOES data dating back decades are publicly available to scientists.

24. Are the data archived at another facility or are there plans to do so? Please explain.

No

25. Is there an existing agreement or requirement driving this request to archive? Have you already contacted someone at NCEI?

NCEI is responsible for archiving L2+ data and has received funding to do so (see below).

26. Do you have a data management plan for your data?

No

27. Have funds been allocated to archive the data at NCEI?

\$411K/year is available through SPADES-IDP Operations and Maintenance (O&M) funding. Much of this money will go to tasks such as transitioning the demonstration L2+ software system to operations and updating the demonstration system on an ongoing basis. However, some of the money is allocated for archive purposes.

L2+ Archive Break-out funding profile:

FY16 (\$25K) // FY17 (\$25K) // FY18-36 (\$50K)

28. Identify the affiliated research project, its sponsor, and any project/grant ID as applicable.

N/A

29. Is there a desired deadline for NCEI to archive and provide access to the data?

Archive by: 2016-10-13

Accessible by: 2017-04-13

30. Add any other pertinent information for this request.

In the event that the NWS operational system to produce L2+ data is not functioning at the time space weather instruments are turned on, NCEI will archive the demonstration-system L2+ products. Once the operational data become available, these will be archived instead.

The figure of 68 files per day was arrived at as follows: there are 34 L2+ algorithms total, and each satellite's products will be packaged separately. Files will be tarred by algorithm type. So when two GOES-R series satellites are operational, there will be 68 files/day.

File volumes depend heavily on SUVI image sizes. It is possible that these image files could be archived in a compressed format, which would reduce overall volumes. Since compression format is currently unknown, this RTA includes volumes for uncompressed files.